

# Detection and Evaluation of Physical Therapy Exercises by Dynamic Time Warping Using Wearable Motion Sensor Units

Aras Yurtman and Billur Barshan

**Abstract** We develop an autonomous system that detects and evaluates physical therapy exercises using wearable motion sensors. We propose an algorithm that detects all the occurrences of one or more template signals (representing exercise movements) in a long signal acquired during a physical therapy session. In matching the signals, the algorithm allows some distortion in time, based on dynamic time warping (DTW). The algorithm classifies the executions in one of the exercises and evaluates them as correct/incorrect, giving the error type if there is any. It also provides a quantitative measure of similarity between each matched execution and its template. To evaluate the performance of the algorithm in physical therapy, a dataset consisting of one template execution and ten test executions of each of the three execution types of eight exercises performed by five subjects is recorded, having a total of 120 and 1,200 exercise executions in the training and test sets, respectively, as well as many idle time intervals in the test signals. The proposed algorithm detects 1,125 executions in the whole test set. 8.58 % of the 1,200 executions are missed and 4.91 % of the idle time intervals are incorrectly detected as executions. The accuracy is 93.46 % only for exercise classification and 88.65 % for simultaneous exercise and execution type classification. The proposed system may be used for both estimating the intensity of the physical therapy session and evaluating the executions to provide feedback to the patient and the specialist.

---

A. Yurtman (✉) · B. Barshan

Department of Electrical and Electronics Engineering, Bilkent University, TR-06800

Ankara, Bilkent, Turkey

e-mail: yurtman@ee.bilkent.edu.tr

B. Barshan

e-mail: billur@ee.bilkent.edu.tr

## 1 Introduction

Physical therapy is a widely used type of rehabilitation in the treatment of patients with various disorders. The patients mostly need to repeat one or more exercise movements advised by the specialist in physical therapy sessions. In hospitals or rehabilitation centers, specialists monitor the patients and provide feedback about their performance. However, they often alternate between the patients and cannot monitor each patient continuously [2]. They cannot count the number of correct executions for each patient, and hence cannot estimate the intensity of the therapy session. They provide subjective and qualitative feedback. Moreover, once they learn how to perform them, most patients continue their physiotherapy exercises at home with no feedback at all; hence, they are likely to execute the exercises erroneously [13]. For this purpose, an autonomous system is developed to detect and evaluate physical therapy exercises using wearable motion sensor units.

## 2 Related Work

Several different sensor modalities are used in physical rehabilitation, including inertial, visual, strain, and medical sensors. However, many of the earlier studies are based on estimating the activity/therapy intensity [1, 6, 10] or the energy expenditure [8] using the sensors rather than determining the accuracy of the physical therapy exercises. In numerous studies, a 3-D real-time human body model is constructed to observe the movements [1, 2, 4, 11, 12, 18, 19]. However, in the previous studies, either the exercise executions are cropped manually, the subject marks each execution by pressing a button, or the subject performs each execution when s/he is informed by the system by a sound or on-screen notification. Furthermore, no idle time periods are involved in the studies evaluating the executions.

The following studies are most similar to ours:

- In myHeart neurological rehabilitation concept [5], the accuracy of arm movements are determined by applying a threshold to the result of an open-end variant of DTW, based on the signals acquired from a garment containing strain sensors. A different threshold is used for each subject, and a classification accuracy of 85 % is achieved. The system is limited to the arm movements only and wearing the garment may be difficult for some patients.
- In [13], strain sensors worn on the arm are used to provide real-time feedback to neurological patients undergoing motor rehabilitation. Seven exercises are executed by a healthy subject wearing a left-handed sensorized long-sleeve shirt both correctly and incorrectly at various speeds. The system checks whether the measured signals “match *at most once* a prefix of one of several stored references, used as templates” [13] based on the open-end DTW in order to classify and evaluate the executions. The disadvantage of the proposed system is again the difficulty of wearing the sensorized shirt.

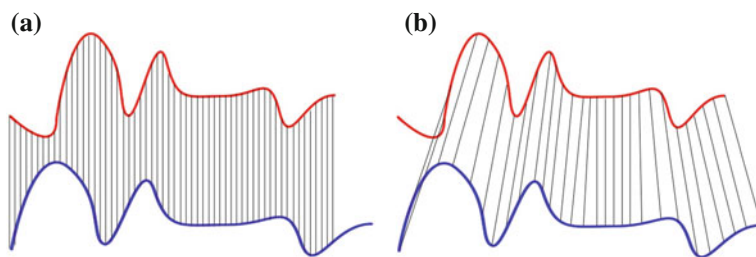
- m-Physio platform [9] classifies the physical rehabilitation activities as correct/incorrect using accelerometers. In m-Physio, the smartphone iPhone containing a tri-axial accelerometer is mounted on the patient's leg or arm depending on the exercise he performs. Some parameters about the exercise durations and filtering of the signals are predetermined. The exercises are classified by some rules as well as by the standard DTW algorithm applied to the accelerometer signals to compare the exercise performed with the pre-recorded template. The limitation of this methodology is the need to determine the parameters for each exercise, the need to touch the iPhone screen to mark the start and end of each exercise execution, and the utilization of the sensors at one location.

Unlike these studies, our system automatically detects the executions in a therapy session and evaluates them while handling idle time periods.

### 3 Multi-Template Multi-Match Dynamic Time Warping

To detect the occurrences of multiple exercise templates in a recorded signal, we propose to use a novel algorithm, namely, *multi-template multi-match dynamic time warping (MTMM-DTW)*, based on dynamic time warping (DTW), both of which are explained below. This method makes it possible to identify correct and incorrect executions of an exercise (identifying the type of error if there is any), the counting of the exercises, and their classification over different exercise types.

**Dynamic time warping** DTW is an algorithm that nonlinearly transforms the time axes of two signals individually such that the transformed signals are most similar to each other, where the (dis)similarity is often measured by the Euclidean distance (see Fig. 1). The minimized distance is called *DTW distance* [7]. DTW can be applied to multi-dimensional signals. The computational complexity of the DTW algorithm is  $\mathcal{O}(NM)$  where  $N$  and  $M$  are the lengths of the two signals.



**Fig. 1** Comparison of the Euclidean and DTW distance measures. **a** The Euclidean distance compares the samples at the same time instants, whereas **b** the DTW distance compares the samples that belong to similar shapes with each other to minimize the distance. Retrieved from [http://upload.wikimedia.org/wikipedia/commons/6/69/Euclidean\\_vs\\_DTW.jpg](http://upload.wikimedia.org/wikipedia/commons/6/69/Euclidean_vs_DTW.jpg)

DTW algorithm can be modified to search for one signal (*template*) inside the other signal (*test signal*) by matching the template with the subsequence of the test signal that minimizes the DTW distance. This algorithm is called *subsequence DTW* [7]. Although the subsequence and the transformations of the time axes of the two signals are jointly optimized, the subsequence DTW algorithm has the same computational complexity as the standard DTW algorithm.

**Multi-template multi-match DTW** We develop the MTMM-DTW algorithm based on subsequence DTW to detect all the occurrences of one or more template signals in a long signal with transformation of the time axes as in DTW. The MTMM-DTW algorithm works iteratively as follows: In each iteration, the subsequence DTW is executed separately for all of the template signals, and the subsequence with the smallest distance is selected as the matching subsequence of that iteration. The matched subsequence of the test signal is made invisible to the future subsequence DTW executions to prevent it from matching again in the future iterations. (10%-prefix and 10%-suffix of the subsequence are left visible to allow some overlapping between the matched subsequences.) If no subsequence is matched in the iteration due to certain constraints (see [16] for details), the MTMM-DTW algorithm stops. In this way, possibly multiple subsequences are obtained, each of which matches to one of the templates with a DTW distance.

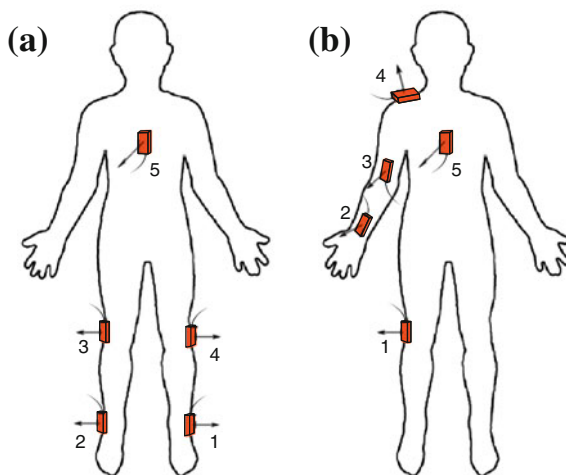
We use the MTMM-DTW algorithm with the templates being different execution types of the exercises to classify the executions in one of the execution types of one of the exercises. By including correct and incorrect execution types of the exercises, the algorithm can be used to evaluate the executions as correct/incorrect, giving the error type if there is any. It can also provide a quantitative measure of similarity between each matched execution and its template.

## 4 Experiments and Results

Five MTx units manufactured by Xsens Technologies [15] are fixed to different positions on the subject's body. Each unit contains three tri-axial devices: an accelerometer, a gyroscope, and a magnetometer. Two different sensor configurations are used to capture leg and arm movements (see Fig. 2).

Eight physical therapy exercises are considered, which are the most commonly assigned exercises to patients, mostly for orthopedic rehabilitation. They were suggested and approved by a physical therapy specialist [14]. The exercises are shown in Fig. 3. Exercises 1–5 and 6–8 use leg and arm sensor configurations, respectively.

In the experiments, there are 5 subjects, 8 exercises, and 3 execution types of each exercise: one correct and two incorrect. For the incorrect execution types, we select the two most common errors for each exercise: fast execution (*type-1 error*) and execution in low amplitude (*type-2 error*). For the templates, each subject performs each execution type of each exercise three times and one of them is selected as the template for that particular execution type, exercise, and subject so that there are  $3 \times 8 \times 5 = 120$  templates in the dataset. For the test dataset, for each subject and exercise, the subject simulates a physical therapy session by executing the exercise



**Fig. 2** The two sensor configurations for **a** the *right leg* and **b** the *right arm* movements. The sensor units are shown as *boxes* with the *arrows* and the *cables* indicating the *z* and *−x* directions, respectively

10 times correctly, waiting idly for some time, then executing the exercise 10 times with type-1 error, waiting idly, and then executing it 10 times with type-2 error. There are a total of 1,200 executions in the test dataset.

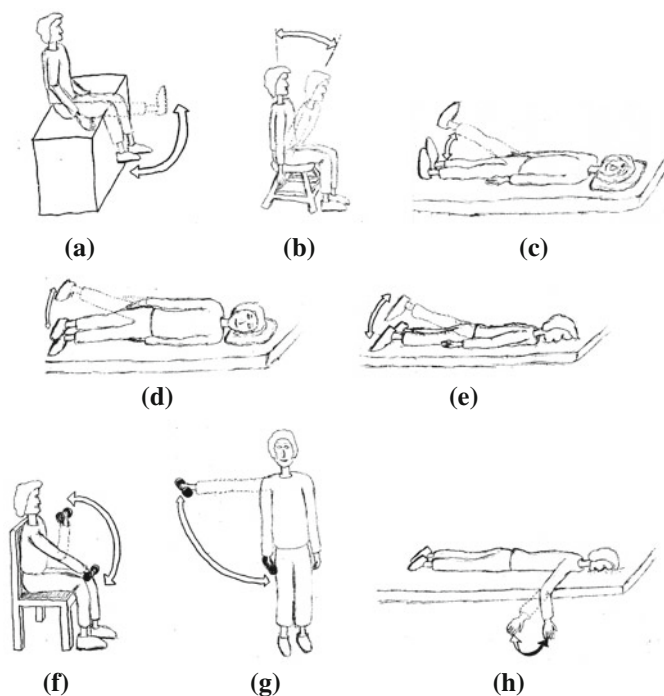
Before applying the MTMM-DTW algorithm, we normalize the signals such that all of the axes of accelerometer, gyroscope, and magnetometer signals have unit variance on average in the whole dataset. We apply MTMM-DTW to the experimental data such that each subsequence must have a duration of at least half of the matching template and the subsequences are allowed to overlap with each other up to 5 % of their durations in the beginning and at the end.

The results are summarized as follows (see Table 1): 1, 125 executions are detected in the whole dataset containing 1,200 executions. This shows that the system makes −6.25 % error in counting the exercises. The average misdetection (MD) and false alarm (FA) rates<sup>1</sup> are 8.58 and 4.91 %, respectively, for the whole dataset. The overall accuracy of the system in exercise classification only<sup>2</sup> is 93.46 %, whereas the overall accuracy in simultaneous exercise and execution type classification<sup>3</sup> is 88.65 %. The overall sensitivity and specificity rates are 91.42 and 95.09 %, respectively, in the whole dataset. The recognized executions and the correctness of their evaluation by the system are shown in Fig. 4 for exercises 1–2 of subject 5. In the whole dataset, there are no exercise misclassifications at all.

<sup>1</sup> The MD and FA rates are calculated as MD rate =  $\frac{\text{\#false negatives}}{\text{\#positives}}$  and FA rate =  $\frac{\text{\#false positives}}{\text{\#negatives}}$ .

<sup>2</sup> The accuracy of exercise classification is calculated as  $\frac{\text{\#correct exercise classifications} + \text{\#true negatives}}{\text{\#positives} + \text{\#negatives}}$ .

<sup>3</sup> The accuracy of exercise and execution type classification is calculated as  $\frac{\text{\#correct exercise and execution type classifications} + \text{\#true negatives}}{\text{\#positives} + \text{\#negatives}}$ .



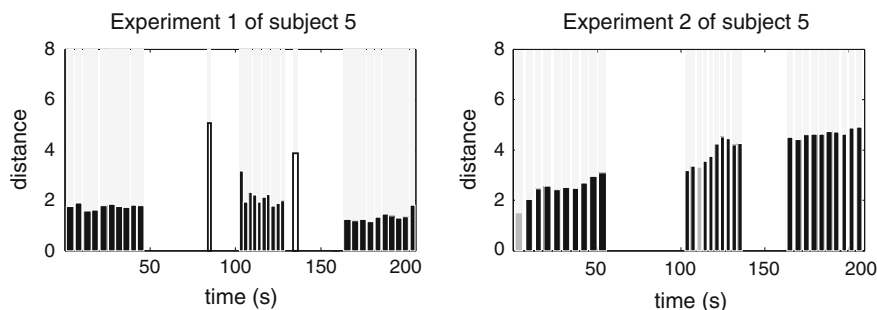
**Fig. 3** Physical therapy exercises. In each exercise, the subject moves his/her leg or arm from the solid position to the dotted position, waits for 5 s, and returns to the solid position. **a** Exercise 1; **b** Exercise 2; **c** Exercise 3; **d** Exercise 4; **e** Exercise 5; **f** Exercise 6; **g** Exercise 7; **h** Exercise 8

**Table 1** Summary of experimental results

Number of detected executions	1,125
Number of actual executions	1,200
Accuracy of exercise classification	93.46 %
Accuracy of exercise and execution type classification	88.65 %
Misdetetection rate	8.58 %
False alarm rate	4.91 %
Sensitivity	91.42 %
Specificity	95.09 %

## 5 Discussion

Detection and classification of the executions are based on prerecorded templates; thus, the patient needs to execute the exercise(s) once for each execution type under the supervision of a physical therapy specialist to record the templates. The execution types contain both correct and incorrect ways of executing a particular exercise. Then, in a physical therapy session, the system automatically detects and provides



**Fig. 4** Detection and classification of exercise executions in exercises 1–2 of subject 5. Each detected execution is shown as a bar whose width is the execution’s duration and height is the DTW distance between the detected subsequence and the matching template. Black and gray filled bars indicate the correct and incorrect execution type classifications, respectively, when the exercise is classified correctly. Unfilled bars indicate the executions classified incorrectly

feedback about the exercise executions. Unlike body model- or rule-based systems, the proposed system requires no adjustments or configuration specific to patients, exercises, sensor types and placements; the only need is recording of the template executions. This eliminates the need for an engineer to reconfigure the system for every new exercise or execution type. Moreover, any sensor configuration representing the movements properly may be directly used without making any adjustments to the algorithm, provided that the configuration is the same in recording template signals and using the system. Furthermore, other sensor types can also be used with the same algorithm. For instance, the wearable inertial sensors and magnetometers can be replaced by the tags of a tag-based localization system that estimates the 3-D positions of the wearable tags and exactly the same methodology can be used for evaluation of the exercise executions [17].

For the detection of multiple occurrences of multiple templates, we select an approach based on DTW because the DTW algorithm is much more flexible than the absolute and Euclidean distance measures in comparing two signals since it tries to match similar parts of the signals. This may be beneficial when the variation in different executions of the same physical therapy exercise is considered. The speeds or durations of some parts of the exercise movement may change, which should be tolerable. For instance, if the exercise contains a phase at which the patient waits for 5 s, the distance should not increase significantly when the patient waits for 4 or 6 s. On the other hand, at the same time, the distance measure should not tolerate significant differences in amplitude that may occur, for example, when the patient waits for 5 s in a different position or orientation. If the absolute or Euclidean distance measures are used, both variation types (amplitude and time) affect the distance value and it is not possible to allow one of them while penalizing the other. In contrast, the DTW algorithm naturally compensates linear or nonlinear changes in the time (or sample) axis but not changes in amplitude, which is desirable in this scheme.

Unlike most of the previous studies, once the exercise movements are recorded, the system proposed in this paper automatically detects the movements as well as the idle time periods, if there are any, during an exercise session, independent of the number of exercise types. The system also classifies each movement as one of the exercise types and evaluates it, indicating the error type if there is any. The patient neither needs to press a button in the beginning and the end of each execution nor select the exercise type that he is going to perform. A physiotherapist is needed only while recording the movements in order to make sure that the patient performs the exercise correctly or with a predetermined error. Then, the patient can perform the exercises anywhere provided that he properly wears the sensors, and can observe how well he performs. Since the system also counts the executions, it can be used to notify the patient when he completes the advised number of repetitions in a given time interval. The results may also be checked by an expert to observe the patient's progress.

## 6 Conclusion and Future Work

In this paper, a novel algorithm, MTMM-DTW, is proposed to detect all of the occurrences of multiple templates in a long signal and is applied to the wearable inertial sensor and magnetometer signals in the area of physical therapy to detect and evaluate individual exercise executions in a physical therapy session. The system is autonomous in that it automatically detects the executions, classifies them as one of the exercises and further classifies them as correctly or incorrectly executed, indicating the error type if there is any. In the newly recorded dataset consisting of 120 templates and 1,200 exercise executions in addition to the idle time intervals, the system makes  $-6.25\%$  error in counting the exercises. The accuracy is  $93.46\%$  for exercise classification and  $88.65\%$  for both exercise and execution type classification. Thus, the system can be used for autonomous feedback in physical therapy.

In a future study, the proposed system may be implemented to run in real-time to provide immediate feedback after each execution. The system can be optimized differently for the individual exercise types and subjects. The experiments may be performed in a more robust way to minimize intra-class variations. Other sensor technologies may be used, such as RF localization [17], to directly get the position information instead of rate information provided by accelerometers and gyroscopes (the second derivative of the linear position and the first derivative of the angular position (angle), respectively). In this case, the relative positions of the RF tags worn on the body can be estimated in 3-D space without any drift errors that exist in inertial sensors. The main drawback of such a system compared to inertial sensing would be that it would radiate radio waves to the environment and would limit the user to a restricted area. Moreover, the system may be used with only one template of the correct execution of each exercise so that the executions are classified in one of the exercises according to the most similar template and then evaluated by applying a threshold to the corresponding DTW distance—if the distance is below the threshold,



meaning that the execution is sufficiently similar to the template, the execution is classified as correct. However, in this case, the threshold would probably have to be determined individually for each exercise or each subject, and it may be difficult to set an appropriate threshold value based on the template signals.

## References

1. Brutovsky J, Novak D (2006) Low-cost motivated rehabilitation system for post-operation exercises. In: Proceedings of 28th annual international IEEE-EMBS. New York, USA, pp 6663–6666 (30 Aug–3 Sept)
2. Fergus P, Kafiyat K, Merabti M, Taleb-bendiab A, El Rhalibi A (2009) Remote physiotherapy treatments using wireless body sensor networks. In: Proceedings of integrated circuit wireless Communication Network. New York, USA, pp 1191–1197 (21–24 June)
3. Keogh EJ, Pazzani MJ (1999) Scaling up dynamic time warping to massive datasets. *Lect Notes Comput Sci* 1704:1–11 (Springer: Berlin, Heidelberg, Germany)
4. Kifayat K, Fergus P, Cooper S, Merabti M (2010) Body area networks for movement analysis in physiotherapy treatments. In: Proceedings of 24th IEEE integrated circuit advanced information network. Perth, Australia, pp 866–872 (20–23 April)
5. Giorgino T, Tormene P, Maggioni G, Pistarini C, Quaglini S (2009) Wireless support to post-stroke rehabilitation: myheart's neurological rehabilitation concept. *IEEE T Inf Technol B*, 13(6):1012–1018 (Nov 2009)
6. Milenkovic M, Jovanov E, Chapman J, Raskovic D, Price J (2002) An accelerometer-based physical rehabilitation system. In: Proceedings of Southeast symposium System. Huntsville, AL, USA, pp 57–60 (18–19 March)
7. Müller M (2007) Information retrieval for music and motion, vol 6. Springer, Berlin, Heidelberg, Germany
8. Pitta F, Troosters T, Probst VS, Spruit MA, Decramer M, Gosselink R (2006) Quantifying physical activity in daily life with questionnaires and motion sensors in COPD. *Eur Respir J* 27(5): 1040–1055
9. Raso I, Hervás R, Bravo J (2010) m-Physio: personalized accelerometer-based physical rehabilitation platform. In: Proceedings of 4th integrated circute mobile ubiquitous computing, systems, services and technologies. Florence, Italy, pp. 416–421 (25–30 Oct)
10. Stéphane C, Mathieu H, Patrick B (2008) Accelerometer-based wireless body area network to estimate intensity of therapy in post-acute rehabilitation. *J Neuroeng Rehabil* 5(1):20–31
11. Tao Y, Hu H (2008) A novel sensing and data fusion system for 3-D arm motion tracking in telerehabilitation. *IEEE IMTC* 57(5):1029–1040
12. Tao Y, Hu H, Zhou H (2007) Integration of vision and inertial sensors for 3D arm motion tracking in home-based rehabilitation. *Int J Robot Res* 26(6):607–624
13. Tormene P, Giorgino T, Quaglini S, Stefanelli M (2009) Matching incomplete time series with dynamic time warping: an algorithm and an application to post-stroke rehabilitation. *Artif Intell Med* 45(1):11–34
14. Tuğcu İ MD (2012) Physical therapy specialist. Department of physical therapy and rehabilitation, Gülhane military medical academy, Turkish Armed Forces Rehabilitation Centre. Personal, communication (June 2012)
15. Xsens Technologies BV (2009) Enschede, The Netherlands. MTi and MTx user manual and technical documentation. <http://www.xsens.com>
16. Yurtman A (2012) Recognition and classification of human activities using wearable sensors. M.S. Thesis, Department of Electrical and Electronics Engineering, Bilkent University, Turkey.

17. Yurtman A, Barshan B (2012) Human activity recognition using tag-based localization (Etiket-tabanlı konumlama ile insan aktivitelerinin tanınması). In: Proceedings of IEEE 20th conference of signal processing communicational Application Fethiye, Muğla, Turkey (18–20 April)
18. Zheng H, Davies RJ, Black ND (2005) Web-based monitoring system for home-based rehabilitation with stroke patients. In: Proceedings of 18th IEEE computational medical system. Dublin, Ireland, pp 419–424 (23–24 June)
19. Zhou H, Hu H (2005) Inertial motion tracking of human arm movements in stroke rehabilitation. In: Proceedings of 2005 IEEE conference mechatronics and automation, vol 3. Ontario, Canada, pp 1306–1311